GENERALIZATION AND ANALYSIS OF THE CONVENTIONAL BEAMFORMER FOR LOCALIZATION OF SPATIALLY DISTRIBUTED SOURCES

Mikael Coldrey (Tapio) and Mats Viberg

Department of Signals & Systems
Chalmers University of Technology, SE-412 96 Göteborg, Sweden
Phone: +46 (0)31 772 1780, Fax: +46 (0)31 772 1748, Email: mito@chalmers.se, viberg@chalmers.se

ABSTRACT

In this paper, we generalize the point source-based conventional beamformer (CBF) to localization of multiple distributed sources that appear in sensor array processing. A distributed source is commonly parameterized by its mean angle and spatial spread. The generalized CBF uses the principal eigenvector of the parameterized signal covariance matrix as its optimal weight vector, which is also shown to be a matched filter. The desired parameter estimates are taken as the peaks of the generalized 2-dimensional beamforming spectrum. Further, the performance of the algorithm is compared numerically to a generalized Capon estimator [1]. Finally, an asymptotic performance analysis of the proposed algorithm is provided and numerically verified.

1. INTRODUCTION

Spatially distributed sources appear in many applications, e.g., sonar, radar, seismology, and wireless communications, to only mention a few. In radio wave propagation the existence of distributed sources has been reported in measurement campaigns such as, e.g., [1].

In wireless communications, it is typically local scattering around the transmitter that makes the transmitted signal to appear having a spatial extension as seen from the elevated base station antenna array. If the signal arrives uncorrelated from several directions, the signal covariance matrix is not rank-one. This leads to that point source-based algorithms deteriorate in performance. This has resulted in numerous publications of new algorithms that can handle full-rank models, e.g., [2, 3, 4, 5, 6], again to only mention a few. Capon's beamformer [7] is a well known point source based algorithm for direction finding which has been generalized to handle distributed sources in [2]. The resulting generalized Capon algorithm was found to outperform the DISPARE [4] and root-MUSIC-based [5] algorithms. In this paper we generalize the conventional beamformer (or Bartlett's beamformer) to also handle distributed sources. The performance of the proposed algorithm is numerically compared to generalized Capon [2], and we provide an asymptotic analysis which is numerically verified.

2. SIGNAL MODELING

In array processing the signals are assumed to be narrowband, which in the presence of p sources yields the following received baseband signal

$$\mathbf{x}(t) = \sum_{i=1}^{p} s_i(t)\mathbf{h}_i + \mathbf{n}(t), \tag{1}$$

The work has been financed in part by the Swedish Foundation for Strategic Research, the Swedish Research Council, and the Personal Computing and Communication Program (PCC++).

where $\mathbf{x}(t)$ is the $K\times 1$ received signal vector (snapshot), $s_i(t)$ and $\mathbf{h}_i(t)$ are the i^{th} source signal and $K\times 1$ channel response vector, respectively. Finally, $\mathbf{n}(t)$ is a $K\times 1$ white noise vector with covariance matrix $\mathbf{R}_n = \mathrm{E}[\mathbf{n}(t)\mathbf{n}^H(t)] = \sigma_n^2\mathbf{I}_K$.

In the point source assumption, the channel response vector for a standard uniform linear array (ULA) is modeled as

$$\mathbf{h}_i = \mathbf{a}(\theta_i) = \left[1, e^{j\pi\sin\theta_i}, \ldots, e^{j(K-1)\pi\sin\theta_i}\right]^T, \tag{2}$$

where θ_i is the direction of arrival (DOA) of the i^{th} source.

A spatially distributed source is commonly described by two parameters, namely its mean angle (DOA) θ_i and spatial spread (standard deviation) σ_{θ_i}. The channel vector of a distributed source can be written as

$$\mathbf{h}_i = \mathbf{h}_i(t;\boldsymbol{\eta}_i) = \int_{\theta\in\Theta_i} \gamma_i(\theta,t;\boldsymbol{\eta}_i)\mathbf{a}(\theta)\,d\theta, \tag{3}$$

where $\boldsymbol{\eta}_i = [\theta_i, \sigma_{\theta_i}]^T$ is the location parameter vector, $\gamma_i(\theta,t;\boldsymbol{\eta}_i)$ represents the underlying spatial amplitude density profile, and Θ_i is the angular support of the i^{th} source, respectively. In this work we consider incoherently distributed (ID) sources [2], although, the proposed algorithm works perfectly well for coherently distributed (CD) sources, and it is shown in [8] that the proposed algorithm is (under certain conditions) asymptotically efficient and coincides with Maximum Likelihood for CD sources. Nonetheless, in the ID case signals arriving from different directions are assumed to be uncorrelated and we have for the i^{th} source

$$\mathrm{E}\left[\gamma_i(\theta,t;\boldsymbol{\eta}_i)\gamma_i^*(\theta',t;\boldsymbol{\eta}_i)\right] = \sigma_{\gamma_i}^2 p_i(\theta;\boldsymbol{\eta}_i)\delta(\theta-\theta'), \tag{4}$$

where $p_i(\theta;\boldsymbol{\eta}_i)$ is the spatial power density function, σ_{γ_i} is the channel gain, and $\delta(\cdot)$ is Dirac's delta-function. Throughout this work, the power density is assumed to be known and it has the same shape but different parameter values for different sources. We also include the channel gain σ_{γ_i} in the source power, i.e., $\sigma_{s_i}^2 \triangleq \sigma_{\gamma_i}^2|s_i(t)|^2$, where $s_i(t)$ is considered being constant modulus and deterministic.

3. CONVENTIONAL BEAMFORMING

In beamforming applications, the DOA of a single deterministic point source is found by directing a spatial filter (beamformer) towards the signal of interest. There are several ways of choosing an optimal beamformer (weight vector $\mathbf{w}_{\text{opt}}$) and the solution depends on the design criterion. Two common beamformers are the conventional (Bartlett) beamformer (CBF) and the Capon beamformer [7], where the former attempts to maximize the expected output power of the spatial filter (which in white noise is equivalent to maximizing the output signal-to-noise ratio (SNR))

$$\max_{\mathbf{w}\ \text{s.t.}\ \mathbf{w}^H\mathbf{w}=1} \mathrm{E}\left[\mathbf{w}^H\mathbf{x}(t)\mathbf{x}^H(t)\mathbf{w}\right]. \tag{5}$$

The maximum is attained when $\mathbf{w}$ is chosen parallel to $\mathbf{a}(\theta)$, i.e.,

$$\mathbf{w}_{\text{opt}} = \frac{\mathbf{a}(\theta)}{\sqrt{\mathbf{a}^H(\theta)\mathbf{a}(\theta)}}. \quad (6)$$

By inserting the optimal weight vector into the expression for the output power of the spatial filter, the DOA estimate is simply found as the peak of the 1-D power azimuth spectrum (PAS)

$$f(\theta) \triangleq \frac{\mathbf{a}^H(\theta)\mathbf{R}_x\mathbf{a}(\theta)}{\mathbf{a}^H(\theta)\mathbf{a}(\theta)}. \quad (7)$$

The DOAs of multiple, say p, point sources are taken as the p largest peaks of the PAS.

4. GENERALIZED BEAMFORMING

The performance of the conventional beamformer for distributed sources was analyzed in [9], where the nominal DOA estimate was taken as the peak of (7). In general, the performance of the conventional peak-finding algorithm is poor and the center of mass of the PAS is a better estimate of the nominal angle of a single distributed source [6].

Here, we reformulate the conventional beamforming problem by taking into account that the signal has a known spatial distribution. Hence, we turn the point source based and non-parametric CBF into a parametric beamforming method, which enables us to find all our parameter estimates.

The expected beamformer output power is given by

$$\mathbf{w}^H\mathbf{R}_x\mathbf{w} = \sigma_s^2\mathbf{w}^H\mathbf{R}_h\mathbf{w} + \sigma_n^2\mathbf{w}^H\mathbf{w}, \quad (8)$$

and is maximized w.r.t. $\mathbf{w}$ s.t. $\mathbf{w}^H\mathbf{w} = 1$. The maximizing weight vector $\mathbf{w}_{\text{opt}}$ is then given by the principal eigenvector of the, generally full-rank, channel vector covariance matrix $\mathbf{R}_h = \text{E}[\mathbf{h}\mathbf{h}^H]$, [10]. The principal eigenvector is defined as the unit norm eigenvector corresponding to the largest eigenvalue $\lambda_{\max}$ of $\mathbf{R}_h$. The resulting beamformer is also optimal in the sense that it maximizes the output SNR, i.e., it is a *matched filter*.

Hence, we can formulate the algorithm as finding the largest peak of the generalized 2-D power spectrum $f(\boldsymbol{\eta}, \mathbf{R}_x)$

$$\hat{\boldsymbol{\eta}} = \arg\max_{\boldsymbol{\eta}} f(\boldsymbol{\eta}, \mathbf{R}_x), \quad \boldsymbol{\eta} = [\theta, \sigma_\theta]^T \quad (9)$$

where $f(\boldsymbol{\eta}, \mathbf{R}_x) \triangleq \mathbf{w}^H\mathbf{R}_x\mathbf{w}$. Here, $\mathbf{R}_x$ contains the true parameters $\boldsymbol{\eta}_0$, and $\mathbf{w} = \mathbf{w}(\boldsymbol{\eta})$ is the principal eigenvector of the matrix $\mathbf{R}_h(\boldsymbol{\eta})$ which is parameterized in the unknown parameters $\boldsymbol{\eta}$. Since $\lambda_{\min} \leq \mathbf{x}^H\mathbf{R}_h\mathbf{x} \leq \lambda_{\max}$ for any unit norm vector $\mathbf{x}$ [10], the maximum value of $f(\boldsymbol{\eta}, \mathbf{R}_x)$ equals $\sigma_s^2\lambda_{\max} + \sigma_n^2$, and is attained by the principal eigenvector $\mathbf{x} = \mathbf{w}$. In general $\lambda_{\max}$ is unique, and the maximum is readily attained by $\hat{\boldsymbol{\eta}} = \boldsymbol{\eta}_0$ when the true covariance matrix is used in (9), which makes the estimator consistent in the sense given below.

In practise we do not have access to the true covariance matrix $\mathbf{R}_x$, so we resort to replacing it by the sample covariance matrix which is a sufficient statistic and is defined as

$$\hat{\mathbf{R}}_x = \frac{1}{N}\sum_{t=1}^{N}\mathbf{x}(t)\mathbf{x}^H(t). \quad (10)$$

It is easily verified that the criterion function $f(\boldsymbol{\eta}, \hat{\mathbf{R}}_x)$ converges uniformly (with increasing numbers of snapshots N) to its limit function $f(\boldsymbol{\eta}, \mathbf{R}_x)$, since

$$\sup_{\boldsymbol{\eta}} \left| f(\boldsymbol{\eta}, \hat{\mathbf{R}}_x) - f(\boldsymbol{\eta}, \mathbf{R}_x) \right| = \sup_{\boldsymbol{\eta}} \left| \mathbf{w}^H(\hat{\mathbf{R}}_x - \mathbf{R}_x)\mathbf{w} \right|$$
$$\leq \left\| \hat{\mathbf{R}}_x - \mathbf{R}_x \right\| \to 0, \text{ w.p.1 as } N \to \infty.$$

Hence, the maximum of $f(\boldsymbol{\eta}, \hat{\mathbf{R}}_x)$ tends to that of $f(\boldsymbol{\eta}, \mathbf{R}_x)$, and the algorithm yields consistent estimates.

As in the point source case, we can identify p distributed sources from the p largest peaks of the generalized 2-D spectrum. In the case of multiple sources, $p > 1$, the estimates will be asymptotically biased due to spectral leakage from closely spaced sources.

Finally, it is also worth noting that in the point source case the channel vector covariance matrix is given by $\mathbf{R}_h(\theta) = \mathbf{a}(\theta)\mathbf{a}^H(\theta)$, which has the principal normalized eigenvector $\mathbf{w} = \mathbf{a}(\theta)/\|\mathbf{a}(\theta)\|$. Thus the generalized 2-D spectrum degenerates to the conventional 1-D spectrum given by (7).

5. NUMERICAL EXAMPLES

In this section we compare the numerical performance of the Generalized conventional beamformer (G-CBF) to that of the generalized Capon (G-Capon) beamformer [2]. In [2] the generalized Capon method was found to outperform the well known DISPARE [4] and root-MUSIC-based [5] algorithms. The Cramér-Rao Lower Bound (CRLB) is also included in the figures.

All estimates are averaged over 2000 independent runs, and the noise is spatio-temporally white with a circular symmetric complex Gaussian distribution.

Example 1:

In the first example we locate a single source with a Gaussian distribution and parameter values $\theta_0 = 0^\circ$ and $\sigma_\theta = 3^\circ$ impinging on a standard uniform linear array (ULA) with $K = 10$ sensors. The SNR is defined as σ_s^2/σ_n^2 and the sample covariance matrix is estimated using $N = 500$ snapshots. In this scenario, G-CBF shows better performance than G-Capon.

Fig. 1 displays the RMSE of the parameter estimates. We observe that the RMSE of the nominal angle estimates attained by the G-CBF is insensitive to the SNR, which is an observation that is consistent with point source estimation using CBF. For low SNR values, the spread estimate of G-CBF improves with increased SNR. For high SNR values, there appears to be a remaining bias (the algorithm is only consistent as $N \to \infty$). The G-Capon spread estimate shows a performance loss for increased SNR, which is likely due to a bias.

Fig. 2 shows the RMSE versus number of sensors. As it is disappointing, it is also interesting to see that the RMSE of the G-CBF spread estimate deteriorates as the number of sensors grows large, and there appears to be an optimal choice (around $K = 10$ sensors) w.r.t. the number of sensors. We will return to this observation in Section 6. The number of sensors has little effect on the G-CBF nominal angle RMSE, and it is only slightly better than the RMSE of the conventional peak-finding algorithm [9], which is included in this figure. There will always be little difference in the nominal angle estimates between the conventional beamformer and its generalized version when the the distributed source can be "seen" as a point source by the array. This is the case when the array beamwidth is large (few sensors) compared to the spread parameter. Hence, it is the relation between the array resolution and the spread parameter that controls how well a distributed source is described by a point source. In this example, the estimates based on G-Capon improves with increasing number of sensors, since we have also increased the number of snapshots according to $N = 20K$. The SNR in this example was 20 dB.

Example 2:

The second example involves locating two Gaussian distributed sources; one with parameter values $\theta_0 = 0^\circ$, $\sigma_{\theta_0} =$

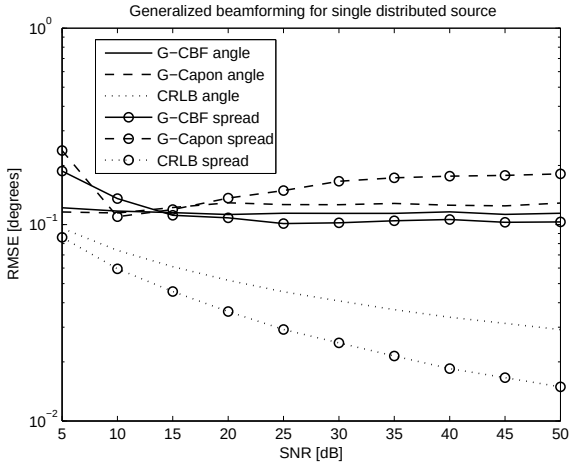

Figure 1: RMSE of $\hat{\theta}_0$ and $\hat{\sigma}_\theta$ for different SNRs. Gaussian distribution with $\theta_0 = 0^\circ$, $\sigma_\theta = 3^\circ$, $K = 10$, and $N = 500$.

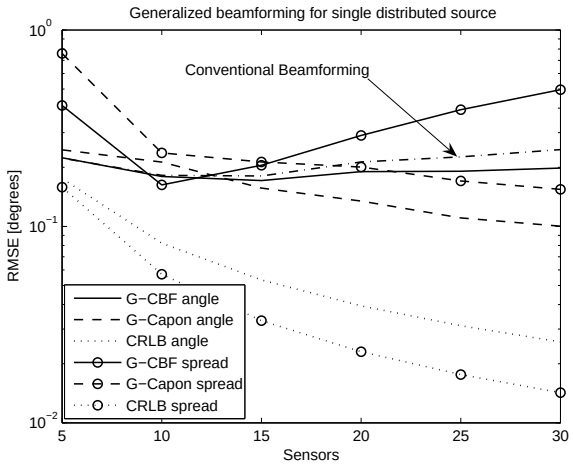

Figure 2: RMSE of $\hat{\theta}_0$ and $\hat{\sigma}_\theta$ versus number of sensors. The figure also includes the angle estimate based on the conventional beamformer [9]. Gaussian distribution with $\theta_0 = 0^\circ$, $\sigma_\theta = 3^\circ$, $N = 20K$, and SNR = 20 dB.

3° and a second with values θ_1 that are varied from 10° to 40°, $\sigma_{\theta_1} = 2^\circ$. A standard ULA with $K = 20$ sensors is used and the SNR is 20 dB per source, i.e., the total received SNR is 23 dB. We see from Fig. 3 that G-CBF captures the nominal angles quite well, while it suffers from spectral leakage when it comes to estimating the spreads of two closely spaced sources.

6. PERFORMANCE ANALYSIS

To simplify the analysis we apply a change of variables and use the concept of spatial frequency and corresponding spread, which are defined as $\omega = 2\pi\Delta \sin\theta$ and $\sigma_\omega = 2\pi\Delta\sigma_\theta \cos\theta$, respectively, where Δ denotes the element separation in wavelengths. Thus, our transformed parameter vector is $\boldsymbol{\psi} = [\omega, \sigma_\omega]^T$. It is well known that the channel covariance matrix $\mathbf{R}_h(\boldsymbol{\psi})$ can be diagonalized as [4]

$$\mathbf{R}_h(\boldsymbol{\psi}) = \mathbf{D}(\omega)\mathbf{B}(\sigma_\omega)\mathbf{D}^H(\omega), \tag{11}$$

where $\mathbf{D}(\omega) = \text{diag}\left(1, e^{j\omega}, \ldots, e^{j(K-1)\omega}\right)$ and $\mathbf{B}(\sigma_\omega)$ is a $K \times K$ Hermitian matrix that depends only on the shape of

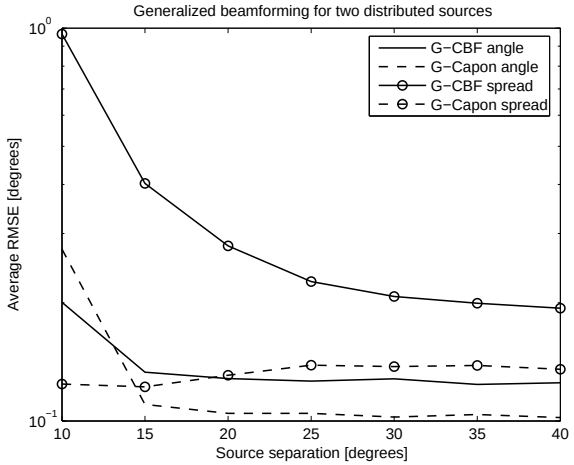

Figure 3: Average RMSE of $\{\hat{\theta}_0, \hat{\theta}_1\}$ and $\{\hat{\sigma}_{\theta_0}, \hat{\sigma}_{\theta_1}\}$ versus source separation. Gaussian distributions with $\{\theta_0 = 0^\circ, \sigma_{\theta_0} = 3^\circ\}$, $\{\theta_1 = 10^\circ \ldots 40^\circ, \sigma_{\theta_1} = 2^\circ\}$, $K = 20$, SNR = 20 dB per source, and $N = 500$.

the angular distribution and is generally of full rank. Closed-form expressions for uniform and Gaussian shapes can be found in, e.g., [11]. For a Gaussian distribution that is symmetric around ω_0 we have

$$[\mathbf{B}(\sigma_\omega)]_{k\ell} = e^{-((k-\ell)\sigma_\omega)^2/2}. \tag{12}$$

It can be shown that $\mathbf{R}_h(\boldsymbol{\psi})$ and $\mathbf{B}(\sigma_\omega)$ share their eigenvalues and that their respective eigenvectors are related according to

$$\mathbf{w}_k(\boldsymbol{\psi}) = \mathbf{D}(\omega)\mathbf{v}_k(\sigma_\omega), \quad k = 1, \ldots, K, \tag{13}$$

where $\mathbf{w}_k$ and $\mathbf{v}_k$ denote the k^{th} eigenvector of $\mathbf{R}_h(\boldsymbol{\psi})$ and $\mathbf{B}(\sigma_\omega)$, respectively. Since the matrix $\mathbf{B}(\sigma_\omega)$ is Hermitian and positive definite, the eigenvalues become real-valued and positive. We also assume that the eigenvalues are simple and sorted in a decreasing order $\lambda_1 > \cdots > \lambda_K > 0$, which provides us with a *unique* (up to a complex scalar) set of orthogonal eigenvectors that are sorted as column vectors into the eigenmatrix $\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_K]$.

Let $f(\boldsymbol{\psi}, \hat{\mathbf{R}}_x) = \mathbf{w}^H(\boldsymbol{\psi})\hat{\mathbf{R}}_x\mathbf{w}(\boldsymbol{\psi})$ denote the criterion function using the sample covariance matrix, and $f(\boldsymbol{\psi}, \mathbf{R}_x) = \mathbf{w}^H(\boldsymbol{\psi})\mathbf{R}_x\mathbf{w}(\boldsymbol{\psi})$ its limit function (as snapshots $N \to \infty$). Applying a standard first order Taylor expansion (neglecting the second and higher order terms) around the true value $\boldsymbol{\psi}_0$ and equating the gradient to zero yields [12]

$$\tilde{\boldsymbol{\psi}} \triangleq \hat{\boldsymbol{\psi}} - \boldsymbol{\psi}_0 \approx -\mathbf{H}^{-1}\nabla f(\boldsymbol{\psi}_0, \hat{\mathbf{R}}_x), \tag{14}$$

where $\mathbf{H}^{-1} \triangleq \left(\nabla^2 f(\boldsymbol{\psi}_0, \mathbf{R}_x)\right)^{-1}$ is the inverse asymptotic Hessian matrix and $\nabla f(\boldsymbol{\psi}_0, \hat{\mathbf{R}}_x)$ is the random gradient, both evaluated at $\boldsymbol{\psi} = \boldsymbol{\psi}_0$. Hence, the covariance matrix of the asymptotic distribution of $\tilde{\boldsymbol{\psi}}$ can be expressed as

$$\text{E}\left[\tilde{\boldsymbol{\psi}}\tilde{\boldsymbol{\psi}}^T\right] \approx \mathbf{H}^{-1}\text{E}\left[\nabla f(\boldsymbol{\psi}_0, \hat{\mathbf{R}}_x)(\nabla f(\boldsymbol{\psi}_0, \hat{\mathbf{R}}_x))^T\right]\mathbf{H}^{-1} \tag{15}$$

To compute the symmetric matrix $\mathbf{H}$, we need the second order partial derivatives of $f(\boldsymbol{\psi}, \mathbf{R}_x)$ w.r.t. ω and σ_ω, which involves finding the first and second order partial derivatives of the principal eigenvector $\mathbf{w}_1(\boldsymbol{\psi})$, since

$$\begin{aligned}[\mathbf{H}]_{ij} &= \frac{\partial^2 f(\boldsymbol{\psi}, \mathbf{R}_x)}{\partial\psi_i\partial\psi_j} \\ &= 2\text{Re}\left\{\mathbf{w}_1^H\mathbf{R}_x\frac{\partial^2\mathbf{w}_1}{\partial\psi_i\partial\psi_j} + \frac{\partial\mathbf{w}_1^H}{\partial\psi_i}\mathbf{R}_x\frac{\partial\mathbf{w}_1}{\partial\psi_j}\right\},\end{aligned} \tag{16}$$

where $i, j = 1, 2$, $\psi_1 = \omega$, and $\psi_2 = \sigma_\omega$. From (13) it follows that the partial derivatives of $\mathbf{w}_1$ simply can be expressed in the derivatives of $\mathbf{D}(\omega)$ and $\mathbf{v}_1(\sigma_\omega)$. Equipped with the previous assumption of simple eigenvalues, the eigenvectors form a orthonormal basis of the whole space $\mathbb{C}^K$. Therefore, we can express the first derivative of *any* eigenvector as a linear combination of the eigenvectors. For the whole eigenmatrix this yields

$$\mathbf{V}'(\sigma_\omega) = \mathbf{V}(\sigma_\omega)\mathbf{C}(\sigma_\omega), \tag{17}$$

for some coefficient matrix $\mathbf{C}$. By taking the derivative w.r.t. σ_ω of the eigensystem $\mathbf{BV} = \mathbf{V}\mathbf{\Lambda}$, where $\mathbf{\Lambda} = \text{diag}(\lambda_1, \ldots, \lambda_K)$ has derivatives $\lambda_k' = \mathbf{v}_k^H \mathbf{B}'(\sigma_\omega)\mathbf{v}_k$, and replacing the eigenmatrix derivative by (17), it can be shown that the off-diagonal entries of the coefficient matrix $\mathbf{C}$ are given by

$$[\mathbf{C}]_{k\ell} = \frac{\mathbf{v}_k^H \mathbf{B}'(\sigma_\omega)\mathbf{v}_\ell}{\lambda_\ell - \lambda_k}, \quad k \neq \ell, \tag{18}$$

where the derivative of $\mathbf{B}(\sigma_\omega)$ is found from (12). The diagonal entries, which represent the component of each eigenvector derivative along the eigenvector itself, can be chosen by introducing some additional normalization. The eigenvectors have unit norm, therefore, by differentiating the inner product $\mathbf{v}_k^H \mathbf{v}_k = 1$, it is readily found that $\text{Re}\{c_{kk}\} = 0$, $k = 1, \ldots, K$. The imaginary part can be found by introducing an additional constraint, such as, e.g., the one that MATLAB® uses in its SVD function where it forces all eigenvectors to have a real-valued first element. To use a less pragmatic approach, we instead use the fact that an eigenvector multiplied by an arbitrary complex scalar also is an eigenvector that has the same corresponding eigenvalue. We can preferably let the complex scalars have unit norm to preserve the unit norm of the eigenvectors. Thus without any loss of generality, we can always choose a new phase rotated orthonormal basis of eigenvectors whose derivative of each eigenvector has no component along the eigenvector itself, i.e., $c_{kk} = 0$, $k = 1, \ldots, K$. The existence of such a phase rotated basis can be shown but the proof is here omitted due to the limited space. This approach is valid as long as the eigenvectors, as here, are functions of one variable.

Once the first derivatives $\{\mathbf{v}_k'(\sigma_\omega)\}_{k=1}^K$ are found it is straightforward to find the second derivative of the principal eigenvector $\mathbf{v}_1''(\sigma_\omega)$ by differentiating the eigensystem twice w.r.t. σ_ω. Once we have the first and second order derivatives of $\mathbf{v}_k$, the partial derivatives of $\mathbf{w}_1$ follow from (13), which inserted into (16) gives the elements of the asymptotic Hessian matrix.

The correlation matrix of the random gradient vector in (15) is found by differentiating $f(\boldsymbol{\psi}, \hat{\mathbf{R}}_x)$ w.r.t. ω and σ_ω in the same fashion as described above. Once the gradient vector is found, a first order perturbation of the gradient vector is performed. By canceling terms that equate to zero and omitting second and higher order terms, the $(i,j)^{\text{th}}$ entry of gradient covariance matrix becomes

$$\begin{aligned}
&\text{E}\left[\nabla f(\boldsymbol{\psi}_0, \hat{\mathbf{R}}_x)(\nabla f(\boldsymbol{\psi}_0, \hat{\mathbf{R}}_x))^T\right]_{ij} \\
&= 2\text{Re}\Big\{\text{E}\Big[\mathbf{w}_1^H \tilde{\mathbf{R}}_x \frac{\partial \mathbf{w}_1}{\partial \psi_i} \mathbf{w}_1^H \tilde{\mathbf{R}}_x \frac{\partial \mathbf{w}_1}{\partial \psi_j} \\
&\quad + \mathbf{w}_1^H \tilde{\mathbf{R}}_x \frac{\partial \mathbf{w}_1}{\partial \psi_i} \frac{\partial \mathbf{w}_1^H}{\partial \psi_j} \tilde{\mathbf{R}}_x \mathbf{w}_1\Big]\Big\},
\end{aligned} \tag{19}$$

where $\tilde{\mathbf{R}}_x$ is the zero mean perturbation matrix, i.e., $\tilde{\mathbf{R}}_x = \hat{\mathbf{R}}_x - \mathbf{R}_x$, and is of order in probability $1/\sqrt{N}$.

Since $N\hat{\mathbf{R}}_x \sim \mathcal{CW}_K(\mathbf{R}_x, N)$, where $\mathcal{CW}_K(\mathbf{R}_x, N)$ denotes a complex central Wishart distribution with N degrees of freedom, we can invoke known expectations of complex Wishart forms. From, e.g., [13] we have

$$\text{E}\left[\mathbf{y}_1^H \tilde{\mathbf{R}}_x \mathbf{y}_2 \mathbf{y}_3^H \tilde{\mathbf{R}}_x \mathbf{y}_4\right] = \frac{1}{N}\left(\mathbf{y}_1^H \mathbf{R}_x \mathbf{y}_4\right)\left(\mathbf{y}_3^H \mathbf{R}_x \mathbf{y}_2\right),$$

where the $\mathbf{y}_i$'s are deterministic vectors. This will yield a closed-form expression for the covariance of the asymptotic distribution. We formulate the final result in the following theorem.

Theorem: Assume $\mathbf{x}(t) \sim \mathcal{N}(0, \mathbf{R}_x(\boldsymbol{\psi}_0))$ *is temporally white and its eigenvalues are single, then the asymptotic distribution of the parameters obtained by inserting the sample covariance matrix into the G-CBF algorithm (9) is given by*

$$\sqrt{N}\left(\hat{\boldsymbol{\psi}} - \boldsymbol{\psi}_0\right) \sim \text{As}\mathcal{N}\left(0, \mathbf{R}_{\tilde{\psi}}\right), \tag{20}$$

where

$$\mathbf{R}_{\tilde{\psi}} = \mathbf{H}^{-1}\mathbf{\Phi}\mathbf{H}^{-1}, \tag{21}$$

$$\begin{aligned}
[\mathbf{\Phi}]_{ij} &= 2\text{Re}\Big\{\Big(\mathbf{w}_1^H \mathbf{R}_x \frac{\partial \mathbf{w}_1}{\partial \psi_j}\Big)\Big(\mathbf{w}_1^H \mathbf{R}_x \frac{\partial \mathbf{w}_1}{\partial \psi_i}\Big) \\
&\quad + \Big(\mathbf{w}_1^H \mathbf{R}_x \mathbf{w}_1\Big)\Big(\frac{\partial \mathbf{w}_1^H}{\partial \psi_j}\mathbf{R}_x \frac{\partial \mathbf{w}_1}{\partial \psi_i}\Big)\Big\},
\end{aligned} \tag{22}$$

$$[\mathbf{H}]_{ij} = 2\text{Re}\Big\{\mathbf{w}_1^H \mathbf{R}_x \frac{\partial^2 \mathbf{w}_1}{\partial \psi_i \psi_j} + \frac{\partial \mathbf{w}_1^H}{\partial \psi_i}\mathbf{R}_x \frac{\partial \mathbf{w}_1}{\partial \psi_j}\Big\}. \tag{23}$$

Proof: A detailed sketch of the proof has already been covered in this section. The missing details can be found in [8].

The asymptotic result in the above theorem holds for the case of single sources. It is easily extended to multiple sources and modeling errors, such as making an erroneous assumption on the true angular distribution, into account. This is easily seen since the case of multiple sources only enter the true covariance matrix $\mathbf{R}_x(\boldsymbol{\psi}_0)$, and the derived asymptotic distribution is a function of this covariance matrix. The case of a model mismatch in the angular distribution is covered by choosing the appropriate $\mathbf{B}(\sigma_\omega)$ matrix. In the mentioned extensions there will, although, be a asymptotic bias present which needs to be taken into account, see [8].

6.1 Validation using Simulated Data

In this section, all estimates are averaged over 2000 independent runs, and the noise is spatio-temporally white with a circular symmetric complex Gaussian distribution.

Example 3:

The purpose of this example is to validate the theoretical expressions for mean and covariance of the asymptotic error distribution. The source has a Gaussian distribution with parameters $\omega_0 = 0$ rad and $\sigma_\omega = 0.1645$ rad. The array is a 10-sensor standard ULA and the SNR is 20dB. The sample covariance matrix is estimated from 500 snapshots and the estimates are averaged over 2000 independent runs. Fig. 4 shows good match between the analytical expressions and the simulated RMSE.

Example 4:

We now return to the earlier and interesting observation that there is an optimal choice of number of sensors that minimize the spread MSE. Table 1 contains the number of antennas that minimize both the analytical and simulated MSE of the spread estimate. Also here the analytical results agree well with simulated results. The optimal choice of number of antennas has a strong dependence on the spatial spread, and

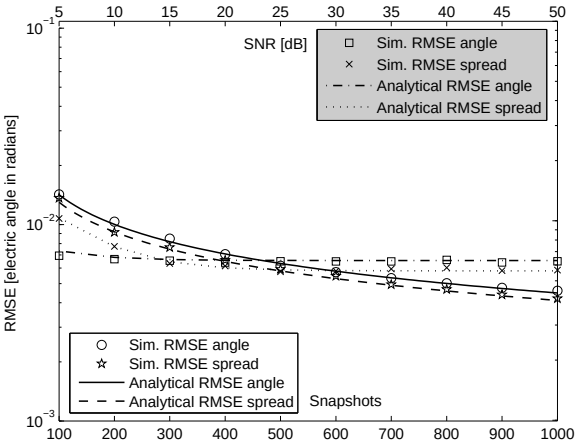

Figure 4: Simulated and analytical RMSE of $\hat{\omega}_0$ and $\hat{\sigma}_\omega$ versus snapshots and SNR. Note the two x-axes. Gaussian distribution with $\omega_0 = 0$ rad, $\sigma_\omega = 0.1645$ rad, $K = 10$, SNR = 20 dB (for the bottom x-axis), and $N = 500$ (for the top x-axis).

σ_θ [deg.]	1	2	3	4	5	6	7	8	9	10
K_{opt} ana.	19	11	8	6	5	4	4	4	3	3
K_{opt} sim.	20	10	7	7	5	4	4	4	3	3

Table 1: Analytical (ana.) and simulated (sim.) optimal number of antennas for different angular spreads. Gaussian distribution with $\theta_0 = 0°$, SNR = 20 dB, and $N = 500$.

both numerical and analytical results indicate that the optimal array should have a Rayleigh beamwidth ($2\pi/K$ in electric angle) approximately equal to the total spatial extension of the source. There is a intuitive explanation to this, which is that the G-CBF makes the best (in terms of maximizing the beamformers' output SNR) rank-one approximation of the distributed source by choosing the principal eigenvector as its optimal weight vector. As long as the Rayleigh beamwidth is equal to or larger than the sources' spatial extension, a rank-one approximation serves as a good approximation, and an increase in the number of antennas will only give better performance. When the number of antennas is increased beyond this limit, the array will experience a source with a larger spatial spread, and the rank-one approximation is no longer a good approximation.

7. CONCLUSIONS

In this paper a generalization of the conventional point source-based beamformer was introduced. The generalized beamformer works as a *matched filter*, and is able to estimate the parameters of multiple spatially distributed sources by locating the largest peaks of the generalized 2-D spectrum. Its performance is compared numerically to a generalized Capon estimator [2] and it shows competitive performance for localization of a single distributed source. Also a performance analysis is provided and the algorithm's interesting "feature" of having an optimal choice w.r.t. the number of antennas is investigated. It turns out that the optimal array that minimizes the MSE is the one that yields a Rayleigh beamwidth approximately equal to the total spatial extension of the source. The intuitive explanation for this is that the G-CBF serves as the best (in terms of maximizing the output SNR) rank-one approximation of the source, and when the number of sensors is increased until the array beamwidth becomes less than the source's spatial extension, the rank-one approximation becomes a poor approximation. This in fact provides an insight to how large the spread values need to be to actually render a source that is experienced as being distributed by the array.

The proposed G-CBF algorithm has a high computational complexity since it for each parameter value in the optimization routine needs to compute a SVD, although for coherently distributed (CD) sources, there is no need to compute the SVD. In fact, the G-CBF algorithm is under certain conditions asymptotically efficient and coincides with Maximum Likelihood for single CD sources, see [8]. Unlike the generalized CBF, the generalized Capon estimator is adaptive (data-dependent). Hence, its analysis takes more elaborate measures and see [8] for a detailed discussion and results regarding this case.

REFERENCES

[1] K.I. Pedersen, P.E. Mogensen, and B.H. Fleury. "A Stochastic Model for the Temporal and Azimuthal Dispersion Seen at the Base Station in Outdoor Propagation Environments". *IEEE Trans. on VT*, 49:437–447, March 2000.

[2] A. Hassanien, S. Shahbazpanahi, and A.B. Gershman. "A Generalized Capon Estimator for Localization of Multiple Spread Sources". *IEEE Trans. on Signal Processing*, 52(1):280–283, January 2004.

[3] T. Trump and B. Ottersten. "Estimation of Nominal Direction of Arrival and Angular Spread using an Array of Sensors". *Signal Processing*, 50(1-2):57–69, Apr. 1996.

[4] Y. Meng, P. Stoica, and K.M. Wong. "Estimation of the Directions of Arrival of Spatially Dispersed Signals in Array Processing". *Proc. IEE Radar, Sonar and Nav.*, 143(1):1–9, Feb. 1996.

[5] M. Bengtsson and B. Ottersten. "Low-Complexity Estimators for Distributed Sources". *IEEE Trans. on Signal Processing*, 48(8):2185–2194, August 2000.

[6] M. (Tapio) Coldrey. "On the use of beamforming for estimation of spatially distributed signals". In *Proc. IEEE ICASSP 03*, volume 5, pages 369–372, Hong Kong, China, April 2003.

[7] J. Capon. "High Resolution Frequency Wave Number Spectrum Analysis". *In Proc. IEEE*, 57:1408–1418, 1969.

[8] M. Coldrey. *"Estimation and Performance Analysis of Wireless Multiple Antenna Communication Channels"*. PhD thesis, Chalmers University of Technology, Göteborg, Sweden, March 2006.

[9] R. Raich, J. Goldberg, and H. Messer. "Bearing Estimation for a Distributed Source via the Conventional Beamformer". *Proc. Stat. Signal Array Process. Workshop*, pages 5–8, Sep. 1998.

[10] P. Stoica and R. Moses. *Introduction to Spectral Analysis*. Prentice Hall, Upper Saddle River, NJ, 1997.

[11] M. Bengtsson. *"Antenna Array Signal Processing for High Rank Data Models"*. PhD thesis, Royal Institute of Technology, Stockholm, Sweden, 1999.

[12] L. Ljung. *System Identification: Theory for the User*. Prentice-Hall, Englewood Cliffs, NJ, 1987.

[13] J.A. Tague and C.I. Caldwell. "Expectations of Useful Complex Wishart Forms". *Multidimensional Systems and Signal Processing*, 5:263–279, 1994.